

Machine Learning Workshop

Digitale Selbstbestimmung Basel – planet13

26. August 2021, überarbeitete Version 1.2

Setup

Wer auf den bereitgestellten Laptops arbeitet, hat bereits das komplette Setup vorbereitet und kann direkt zu den Aufgaben springen.

Ansonsten: Dieser Workshop benötigt das Programm Weka der Universität Waikato und drei Datensets. Davon sind zwei aus dem UCI Machine Learning repository und eines ist selbst generiert. Alle drei Datensets werden von uns zur Verfügung gestellt.

Download Weka

Die grafische Applikation Weka stellt die Algorithmen und einige Analysewerkzeuge zum Trainieren zur Verfügung.

- Website Weka: <https://www.cs.waikato.ac.nz/ml/weka/index.html>
- Download und Installation: https://waikato.github.io/weka-wiki/downloading_weka/
- Die Datei mit Rechtsklick entpacken
- Um Weka zu starten: Die Datei `weka . sh` öffnen

Optional: Download von 37 UCI Datensets

Die folgende Datei enthält 37 öffentliche Datensets aus dem UCI Machine Learning Repository. Diese sind super zum ausprobieren, ihre Beschaffung ist jedoch optional.

Direkter Link: <https://prdownloads.sourceforge.net/weka/datasets-UCI.jar>

`datasets-UCI.jar` mit Rechtsklick entpacken

In der Datei sind 37 verschiedene Datensets enthalten.

Einführung

Für die Aufgaben sollten folgende Punkte aus der Einführung klar sein:

- Was ist ein Datenset mit Reihen und Attributen
- Was ist eine Klassifikationsalgorithmus
- Was ist ein Trainings- und Testset